CRD report

News from the LBNL Computational Research Division — September 2004

LBNL's Evaluation of Earth Simulator Performance Nominated for Best Paper Award at SC2004

With the re-emergence of viable vector computing systems such as the Earth Simulator and the Cray X1, there is renewed debate about which architecture is best suited for running large-scale scientific applications.

In order to cut through the conflicting claims of fastest, biggest, etc., a team led by Lenny Oliker of CRD's Future Technologies Group put five different systems through their paces, running four different scientific applications key to DOE research programs. As part of the effort, the group became the first international team to conduct a performance evaluation study of the 5,120-processor Earth Simulator. The team also assessed the performance of:


Lenny Oliker

- the 6,080-processor Power3 IBM supercomputer running AIX 5.1 at the NERSC Center at Lawrence Berkeley National Laboratory
- the 864-processor Power4 IBM supercomputer running AIX 5.2 at Oak Ridge National Laboratory
- the 256-processor SGI Altix 3000 system running 64-bit Linux at ORNL

(continued on page 2)

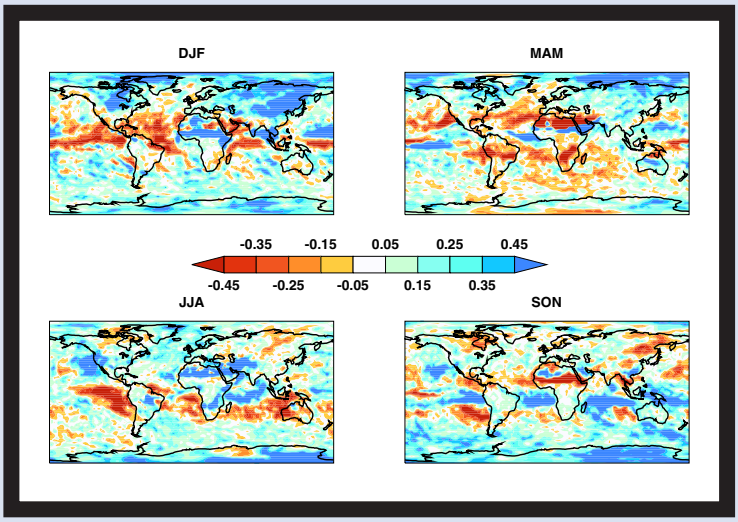
CRD Technical Reports Now Indexed on the Web

CRD webmaster John Hules has created an up-to-date list of publications by members of the division.

The lists are available in several formats, organized chronologically, alphabetically and by department. The overall chronological list can be found at <http://crd.lbl.gov/html/news/techreports.html>.

Michael Wehner to Analyze Climate Models for International Report on Climate Change

Michael Wehner, a climate researcher in CRD's Scientific Computing Group, has received a National Science Foundation grant to analyze the results of three new climate models as a means of determining their predictive quality. Each of the three models will be run to predict both past and future climate change patterns. The results will also be compared with observational climate data to see how the predictions and observations correlate.



These images show the fractional change from 1979 to 2079 in the 20-year return value of seasonal maxima of daily precipitation as computed using the Parallel Climate Model.

"Some features will be correct, some will be wrong and some will be difficult to determine," said Wehner. "Our goal is to learn about the ability of the models to predict reality, and hopefully we will learn more

(continued on page 3)

ACTS Collection Sponsors Fifth Annual Workshop

Since the first workshop aimed at familiarizing users with the DOE Advanced CompuTational Software (ACTS) Collection was held five years ago, organizers Tony Drummond and Osni Marques have seen a real shift in the participants.

Initially, the 50 or so attendees seemed new to many aspects of HPC, but were eager to learn about the ACTS Collection, a set of software tools aimed at simplifying the solution of common and important computational problems. Starting last year and continuing in 2004, however, the participants arrived ready to immediately start working with the tools, Drummond and Marques said.

"This year, we noticed that a lot of the people were really ready to use the tools, which is a change from the first years," said Drummond. "They asked more questions and even took advantage of their two-week post-workshop access to the Seaborg supercomputer at NERSC."

The 2004 workshop, held Aug. 24–27 at LBNL, drew 45 participants, including students, researchers and industry representatives. Drummond and Marques, who are members of CRD's Scientific Computing Group, organize the workshop in their role as managers of the ACTS Collection.

The workshop featured speakers from LBNL, as well as Argonne, Lawrence Livermore, Oak Ridge and Pacific Northwest national labs, and Sandia National Laboratories. The workshop also featured hands-on sessions to give attendees experiences in using the tools.

(continued on page 3)

Evaluation of Earth Simulator, Other Architectures *(continued from page 1)*

- the 512-processor Cray X1 supercomputer running UNICOS at ORNL.

The results of the comparison are of great interest to the HPC community. The team's paper was accepted for the SC2004 conference, then nominated for Best Paper. The winning paper will be announced at the conference in November.

Earth Simulator

In addition to Oliker, the team includes Julian Borrill, Andrew Canning, Jonathan Carter and John Shalf, all of LBNL, and Stephane Ethier of Princeton Plasma Physics Laboratory.

NERSC's IBM supercomputer, seaborg

"This effort relates to the fact that the gap between peak and actual performance for scientific codes keeps growing," said Oliker, who won the Best Paper Award at SC99. "Because of the increasing cost and complexity of HPC systems, it is critical to determine which classes of applications are best suited for a given architecture."

In their abstract, the group members write, "Computational scientists have seen a frustrating trend of stagnating application performance despite dramatic increases in the claimed peak capability of high performance computing systems. This trend has been widely attributed to the use of superscalar-based commodity components whose architectural designs offer a balance between memory performance, network capability, and execution rate that is poorly matched to the requirements of large-scale numerical computations."

The four applications and research areas selected by the team for the evaluation are:

- Cactus, an astrophysics code that evolves Einstein's equations from the Theory of Relativity using the Arnowitt-Deser-Misner method
- GTC, a magnetic fusion application that uses the particle-in-cell approach to solve nonlinear gyrophase-averaged Vlasov-Poisson equations
- LBMHD, a plasma physics application that uses the Lattice-Boltzmann method to study magnetohydrodynamics
- PARATEC, a first-principles materials science code that solves the Kohn-Sham equations of density-functional theory to obtain electronic wave functions.

So, what are the team's conclusions?

"The four applications successfully ran on the Earth Simulator with high parallel efficiency," Oliker said. "And they ran faster than on any other measured architecture — generally by a large margin."

However, Oliker added, only codes that scale well and are suited to the vector architecture may be run on the Earth Simulator.

"Vector architectures are extremely powerful for the set of applications that map well to those architectures," Oliker said. "But if even a small part of the code is not vectorized, overall performance degrades rapidly."

As with most scientific inquiries, the ultimate solution to the problem is neither simple nor straightforward.

"We're at a point where no single architecture is well suited to the full spectrum of scientific applications," Oliker said. "One size does not fit all, so we need a range of systems. It's conceivable that future supercomputers would have heterogeneous architectures within a single system, with different sections of a code running on different components."

The team's full paper can be found at <http://crd.lbl.gov/~oliker/papers/SC04.pdf>.

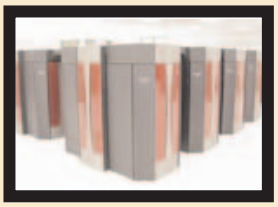
Cosmology Code's Performance Characteristics Lead to Paper

When the team from LBNL's CRD and NERSC divisions spent a week conducting performance evaluations of Japan's Earth Simulator in late 2003, one of the five codes they intended to run did not scale well enough to be used.

According to team leader Lenny Oliker, the application MADCAP (the Microwave Anisotropy Dataset Computational Analysis Package), a parallel implementation of cosmic microwave bacground map-making and power spectrum estimation algorithms, has been tuned since then and the group is returning to Japan in October 2004 to try again. MADCAP's performance was hampered by the lack of a fast global file system on the Earth Simulator, since the code has high I/O requirements.

The results of the MADCAP evaluation runs are the basis for the paper "Performance Characteristics of a Cosmology Package on Leading HPC Architectures," which will be presented at HiPC 2004, the eleventh International Conference on HPC to be held in Bangalore, India, in December. The paper is available at <http://crd.lbl.gov/~oliker/papers/HIPC04.pdf>.

ORNL's Altix



ORNL's CRAY X1

ORNL's IBM Power 4

Climate Models *(continued from page 1)*

about the reality of climate change in the process."

When finished, his work will be incorporated into an international report on climate change. The report is designed to advise governments on the state of scientific research on climate change so that polices can be set.

The first part of Wehner's research will build on his fruitful collaboration with Ben Santer of the Program for Climate Model Diagnosis and Intercomparison (PCMDI) at the Lawrence Livermore National Laboratory. In a series of studies, the team led by Santer has used the Parallel Climate Model (PCM) and observational data to show that statistically significant changes to the tropospheric temperature, lower stratospheric temperature, and tropopause height have occurred in the recent past, and that these changes were the result of human activities. This work was reported in the July 25, 2003, edition of Science magazine.

Wehner will extend a multivariate climate change detection analysis, which he and colleagues have already applied to the older PCM, to early simulations from the three most recent U.S. coupled climate models: the Goddard Institute for Space Studies (GISS) model, the Geophysical Fluid Dynamics Laboratory (GFDL) model, and the Community Climate System Model (CCSM3) from the National Center for Atmospheric Research.

The second part of Wehner's research is to characterize the ability of coupled climate models to simulate extreme weather events. While the mathematical formalisms necessary for rigorous statistical description of extreme events are quite well developed, application to climate models is still in its infancy.

The ability of these models to accurately simulate historical climate changes in the twentieth century will determine the credibility of their predictions of climate change during the twenty-first century and beyond.

"Five years ago, the results from the climate change models were at odds with the observational data," Wehner said. "But now, with our ability to simulate sea, ice and land systems, we are finding real similarities between the observational data, particularly from satellites, and the models."

Wehner's proposal, titled "Multivariate Climate Change Detection," was one of several funded as part of the U.S. CLIVAR (Climate Variability and Predictability) project's Climate Model Evaluation Project (CMEP). The objective of CMEP is to increase diagnostic research into the quality of model simulations, leading to more robust evaluations of model predictions and a better quantification of uncertainty in projections of future climate. U.S. CLIVAR is a joint project of NSF, NOAA, NASA and DOE, and is part of an international, interdisciplinary research effort within the World Climate Research Programme.

The first step is getting the necessary climate data in place. Wehner is currently transporting data files from Oak Ridge National Lab and the National Center for Atmospheric Research. Wehner will then run the models using the data.

The results of the CMEP analyses will be presented at a CLIVAR workshop in Hawaii from March 1–4, 2005. A workshop report summarizing the presentations will be furnished to the lead authors of the relevant chapters for the Intergovernmental Panel on Climate Change's (IPCC) Assessment Report Four, a comprehensive review of current knowledge about climate change, which will be published in 2007.

"Until now, most of the science for the IPCC has been done by Europeans and the time is right for Americans to begin making a major impact in this field by becoming a major source of the science, not just analyzing the European data," Wehner said. "By showing that our models, such as the CCSM3, are credible, we are poised to make a major contribution. This is really a high point in my career."

ACTS Collection Sponsors Workshop *(continued from page 1)*

The workshop facilitates one-on-one exchanges between tool developers and users. Additionally, this year's event drew industry participation ranging from major players in HPC technology (HP and IBM) to commercial library and application software development companies (IMSL). All industrial participants contributed feedback and complimented the ACTS Collection project and the individual tool development projects.

"Increasing the visibility of the ACTS collection among industrial HPC collaborations helps guarantee the long-term support for the robust software we develop and support under ACTS," Drummond added

More information about the workshop can be found at <http://acts.nersc.gov/events/Workshop2004/>. Workshop materials can be found at <http://acts.nersc.gov/events/Workshop2004/slides>.

What is CRD Report?

This is CRD Report, a publication highlighting recent achievements by staff members in Berkeley Lab's Computational Research Division. Distributed every other month via email and posted on the Web at http://crd.lbl.gov/DOEresources, CRD Report may be freely distributed. CRD Report is edited by Jon Bashor, JBashor@lbl.gov or 510-486-5849.